



US009665633B2

(12) **United States Patent**
Dageville et al.

(10) **Patent No.:** **US 9,665,633 B2**
(45) **Date of Patent:** **May 30, 2017**

(54) **DATA MANAGEMENT SYSTEMS AND METHODS**

(2013.01); *G06F 17/30466* (2013.01); *G06F 17/30498* (2013.01); *G06F 17/30545* (2013.01); *G06F 17/30598* (2013.01); *G06F 17/30864* (2013.01); *G06F 17/30867* (2013.01);

(71) Applicant: **Snowflake Computing Inc.**, San Mateo, CA (US)

(Continued)

(72) Inventors: **Benoit Dageville**, Foster City, CA (US); **Thierry Cruanes**, San Mateo, CA (US); **Marcin Zukowski**, San Mateo, CA (US)

(58) **Field of Classification Search**
None

See application file for complete search history.

(73) Assignee: **Snowflake Computing, Inc.**, San Mateo, CA (US)

(56) **References Cited**

U.S. PATENT DOCUMENTS

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

5,787,466 A 7/1998 Berliner
6,490,590 B1 12/2002 Fink
(Continued)

(21) Appl. No.: **14/518,873**

FOREIGN PATENT DOCUMENTS

(22) Filed: **Oct. 20, 2014**

CN 102496060 6/2012
CN 203261358 10/2013

(Continued)

(65) **Prior Publication Data**

US 2015/0234688 A1 Aug. 20, 2015

Primary Examiner — Gregory A Kessler

(74) *Attorney, Agent, or Firm* — David R. Stevens;
Stevens Law Group

Related U.S. Application Data

(60) Provisional application No. 61/941,986, filed on Feb. 19, 2014.

(57) **ABSTRACT**

(51) **Int. Cl.**

G06F 9/46 (2006.01)
G06F 17/30 (2006.01)
G06F 9/50 (2006.01)
G06F 9/48 (2006.01)
H04L 29/08 (2006.01)

Example data management systems and methods are described. In one implementation, a method identifies multiple files to process based on a received query and identifies multiple execution nodes available to process the multiple files. The method initially creates multiple scansets, each including a portion of the multiple files, and assigns each scanset to one of the execution nodes based on a file assignment model. The multiple scansets are processed by the multiple execution nodes. If the method determines that a particular execution node has finished processing all files in its assigned scanset, an unprocessed file is reassigned from another execution node to the particular execution node.

(52) **U.S. Cl.**

CPC *G06F 17/30575* (2013.01); *G06F 9/4881* (2013.01); *G06F 9/5016* (2013.01); *G06F 9/5088* (2013.01); *G06F 17/302* (2013.01); *G06F 17/3048* (2013.01); *G06F 17/30292* (2013.01); *G06F 17/30315* (2013.01); *G06F 17/30371* (2013.01); *G06F 17/30463*

24 Claims, 14 Drawing Sheets

